OLLSCOIL NA GAILLIMHE
UNIVERSITY OF GALWAY

# ***Autumn Examinations 2022-2023***

| | |
|---|---|
| **Course Instance Code(s)** | 4BCT, 1CSD1, 1CSD2, 1OA2, 1MAI1 |
| **Exam(s)** | B.Sc. Computer Science and IT, M.Sc. in Data Analytics, M.Sc. in Artificial Intelligence |
| **Module Code(s)** | CT4100 |
| **Module(s)** | Information Retrieval |
| Paper No. | 1 |
| External Examiner(s) | Dr. R. Trestian |
| Internal Examiner(s) | Professor M. Madden<br>*Dr. C. O'Riordan |

**Instructions:** Answer any 3 questions. All questions are equally weighted.

| | |
|---|---|
| **Duration** | 2 hours |
| **No. of Pages** | 3 |
| **Discipline(s)** | Computer Science |
| **Course Co-ordinator(s)** | Dr. C. O'Riordan, Dr M. Nickles, Dr. F. Glavin |

**Requirements:**

| | |
|---|---|
| Release in Exam Venue | Yes |
| Handout | None |
| Statistical/ Log Tables | None |
| Cambridge Tables | None |
| Graph Paper | None |
| Log Graph Paper | None |
| Other Materials | None |

**Question 1**

(a) Describe, in your own words, with reference to any well-known term weighting scheme, the main constituents of a good weighting scheme. (8)

(b) Discuss the advantages of the axiomatic approaches that have been adopted to design weighting schemes in information retrieval. (8)

(c) Outline a suitable approach to indexing a document collection to allow efficient handling of queries in a system adopting a vector space framework. (9)

**Question 2**

(a) Feedback mechanisms have been adopted to achieve better representations of the user's information need. Discuss a suitable explicit feedback mechanisms that could be used with the vector space model. (8)

(b) Pseudo-feedback mechanisms have been used to augment queries or to suggest candidate terms to users. Suggest an approach to provide a diverse set of candidate terms to users to add to their queries. (8)

(c) Given a submitted query, we can process the query in many ways e.g. query expansion. These additional techniques may be more beneficial for *difficult* queries. Suggest a suitable means to identify a *difficult* query. (9)

**Question 3**

(a) Precision and recall are often used to measure the quality of an answer set. Describe these concepts and explain how you could create a precision-recall graph. (8)

(b) Given a clustering algorithm, discuss a suitable approach to evaluating the quality of the algorithm. (8)

(c) With reference to any existing clustering approach, discuss how these clustering algorithms can be useful in the development of information retrieval systems. (9)

**Question 4**

(a) Describe the main steps taken in collaborative filtering (recommender systems) to produce recommendations for a user. (8)

(b) Consider a music recommendation system, suggest an approach where content information (e.g for each track, the name of the artist, album, year, genre etc.) can be used to improve the coverage of the recommendation system. (7)

(c) With reference to any two information retrieval tasks, outline learning approaches that can be adopted to tackle these tasks. Outline the data available, the learning mechanism and any limitations of the approaches. (10)

**END**